Visualizing Nucleic Acids in a Virus by X-ray Crystallography and Electron Cryomicroscopy

Pariacoto virus (PaV), a single-stranded RNA virus isolated from the Southern armyworm, is the subject of this investigation. In this study, we have used combined x-ray crystallography at the Advanced Photon Source with electron cryomicroscopy (cryoEM) and three-dimensional reconstruction to obtain significant insights into PaV virus structure and assembly at a resolution of 3.0 Å. The stucture of PaV reveals a dodecahedral cage of duplex RNA. The crystal structure and cryoEM reconstruction of PaV have significantly expanded our understanding of the assembly and stability of virus particles, and have made it possible to build an atomic model for the whole viral genome packaged in the capsid and to predict steps and intermediates in the assembly pathway.

Interactions between nucleic acids and proteins play critical roles throughout the life cycle of a virus. The proper and efficient assembly of a virus requires recognition and structural interactions between the viral genome and coat proteins that allow the viral nucleic acids to be packaged into a highly ordered protein shell to form a stable and infective virion. Meanwhile, virus infection requires the viral nucleic acids to be released through the coat protein shell into the host cells for synthesis of viral nucleic acids and proteins and assembly of new virus particles.

The combination of x-ray crystallography with electron cryomicroscopy (cryoEM) and three-dimensional reconstruction has provided significant insights into virus structure and assembly both at atomic detail and at moderate resolution. In most cases, the coat proteins assemble into a well-defined icosahedral lattice, which is routinely employed in virus structure determination to yield an atomic model for the capsid. However, the organization of viral nucleic acids and the interactions with proteins are largely dynamic, and the lack of icosahedral symmetry generates special difficulties for visualizing the nucleic acid in viruses.

Pariacoto virus (PaV) is a single-stranded RNA virus isolated from the Southern armyworm. It is a nonenveloped $T = 3$ icosahedral insect virus belonging to the family *Nodaviridae*. The viral genome consists of two messenger-sense RNA molecules: RNA1 (3011 nucleotides), which encodes the RNA-dependent RNA polymerase, and RNA2 (1311 nucleotides), which encodes protein alpha, the precursor to the viral coat proteins [1]. The PaV particle is an assemblage of 180 copies of coat proteins that co-encapsidate the two single-stranded genomic RNA molecules. Although the protein subunits are chemically identical, they occupy three quasi-equivalent locations (A, B, and C) within the icosahedral asymmetric unit (Fig. 1a). The A subunits form the 12 pentamers of the capsid, whereas the B and C subunits alternate around the icosahedral 3-fold axes to form 20 hexamers. The simplicity of PaV makes it an excellent model system to investigate the structure, assembly, and infection of viruses.

The crystal structure of PaV was determined at 3.0-Å resolution by molecular replacement and real-space averaging. The atomic model of the icosahedral asymmetric unit contains amino acid residues 7–378 and 394–401 of subunit A, 49–382 of subunit B, 51–382 of subunit C, a 25-nucleotide RNA, and numerous solvent molecules (Fig. 1b).

A 25-nucleotide, A-type RNA duplex was

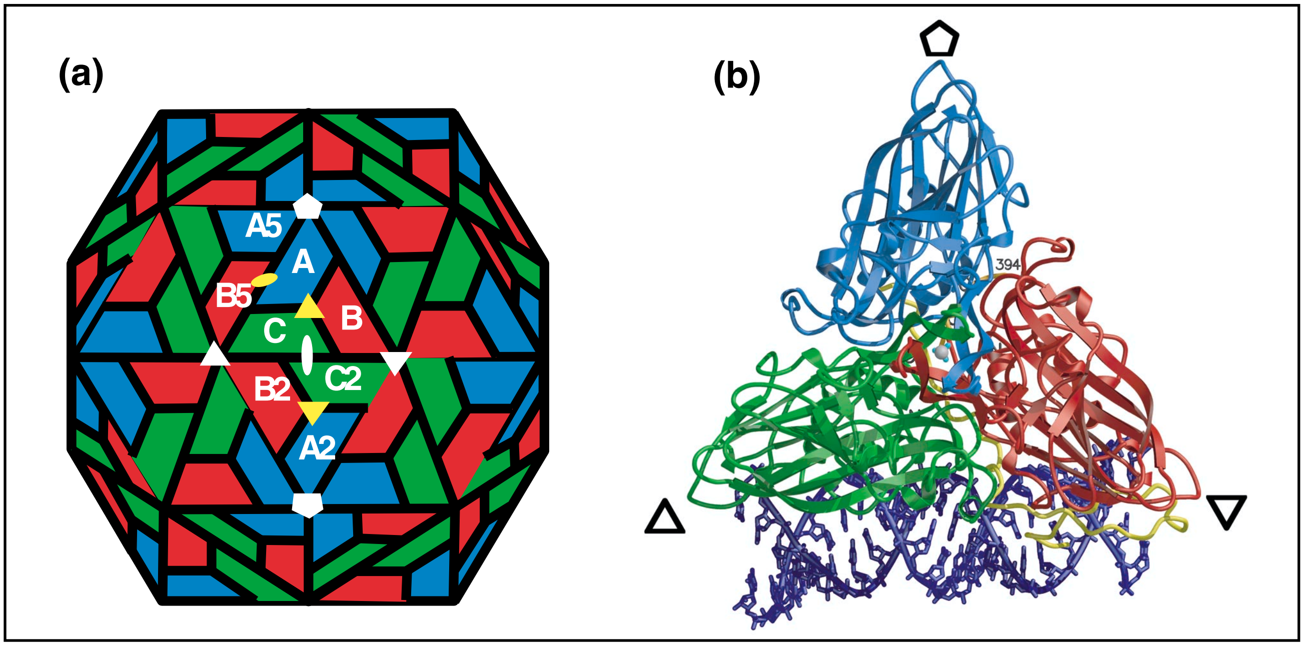

FIG. 1. (a) Schematic presentation of a T = 3 quasi-equivalent lattice corresponding to a rhombic triacontahedron. Each trapezoid represents a single subunit (labeled A, B, and C). The positions of the icosahedral 2-, 3-, and 5-fold axes are indicated by the white ovals, triangles, and pentagons, respectively (yellow for local symmetry). A5: subunit A related by the icosahedral 5-fold symmetry. (b) Overall structure of an icosahedral asymmetric unit of PaV (viewed from outside). Subunits A, B, and C are blue, red, and green, respectively. The extended N-terminus of subunit A is yellow and its C-terminus is gold (residue 394 is labeled). The RNA duplex is shown as a stick model with coils through phosphorus atoms. Ca^{2+} is gray and the sulfate ion (which is mostly hidden behind the Ca^{2+} ion) is cyan. The approximate positions of the 3- and 5-fold axes are indicated by triangles and a pentagon. The same color code is used hereafter unless otherwise stated. Figures were generated with Molscript [2], Bobscript [3], Raster3D [4] and GRASP[5].

modeled into the PaV electron density (Fig. 2). Density comparable to that of well-ordered proteins was visible for the phosphate, sugar, and bases of ribonucleotides 13–17, which lie closest to the icosahedral 2-fold axes. Other ribonucleotides showed weaker but continuous density, consistent with their high-temperature factors in the refined model. Nucleotides 14–17 were modeled as adenine and all others as uridine. Since the structure is averaged by icosahedral symmetry, the 25 ribonucleotides represent an average of the true sequences that appear in each of the 30 unique segments. Nucleotides 13 and 14 of one strand form Watson-Crick base pairs with nucleotides 14 and 13 of the other strand. The base conformation of other nucleotides appears flexible as they do not form strict Watson-Crick base pairs. There is an overhang at the 5' end of each strand,

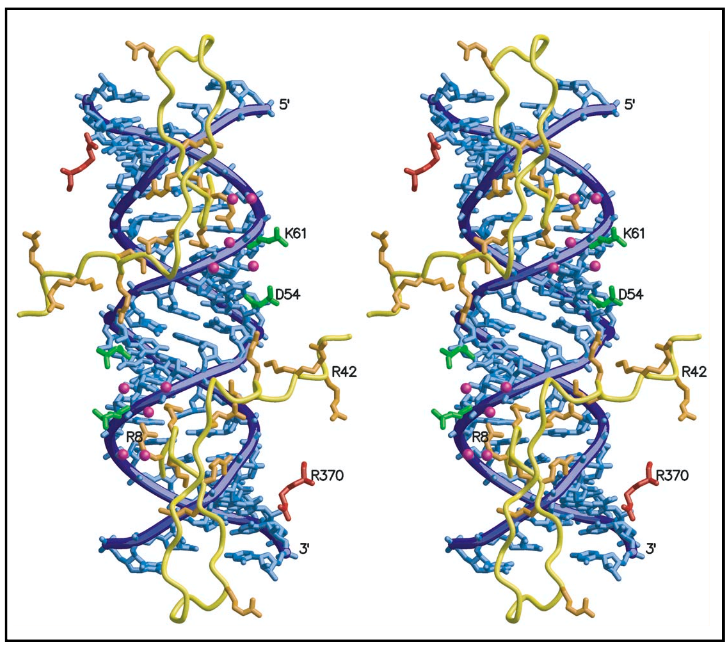

FIG. 2. Interactions between the RNA duplex (blue) and the N-terminal region of the A subunits (yellow). Basic residues from subunits A, B, and C are orange, red, and green, respectively.

which is located near the icosahedral 3-fold axis. The three duplexes around the 3-fold axis appear to form a three-way junction. As a result, a dodecahedral cage is formed by the 30 pieces of ordered RNA duplex in the virus (Fig. 3). This cage provides an "averaged" image of the organization of the ordered portion of the viral RNA. The weaker density at both ends of the RNA duplex near the 3-fold axes implies that adjacent duplexes are not always directly connected by their strands at the 3-fold axes. Instead, the RNA chains sometimes connect to the bulk RNA, which must lie within the dodecahedral cage but does not follow the icosahedral symmetry. The 30 copies of the RNA duplex account for approximately 35% of the total viral genomic RNA (1500 out of 4322 nucleotides).

A cryoEM three-dimensional reconstruction at 23-Å resolution shows a dodecahedral cage as seen in the x-ray structure. A difference map calculated by subtracting a low-resolution map of the crystal structure without RNA from the reconstruction gives a clear view of the RNA cage, which is virtually superposable with the atomic model of RNA in the crystal structure (Fig. 3). There is weak density near the center and between the capsid and the cage, which might be attributed to the disordered portion of the viral RNA.

As in many other viruses, the N-terminal region of PaV coat protein is rich in basic residues. The crystal structure of PaV shows for the first time extensive interactions between the RNA duplex and the extended N-terminal region of subunit A (Fig. 2). The RNA duplex is situated in a groove formed by 2-fold related coat protein subunits (Fig. 4). A total of 24 basic residues in the N-terminal segments of A subunits from two icosahedral asymmetric units line up at the bottom of the groove to produce a positively charged environment that partially neutralizes the abundant negative charge of the RNA phosphate groups. In addition, the guanidine nitrogen

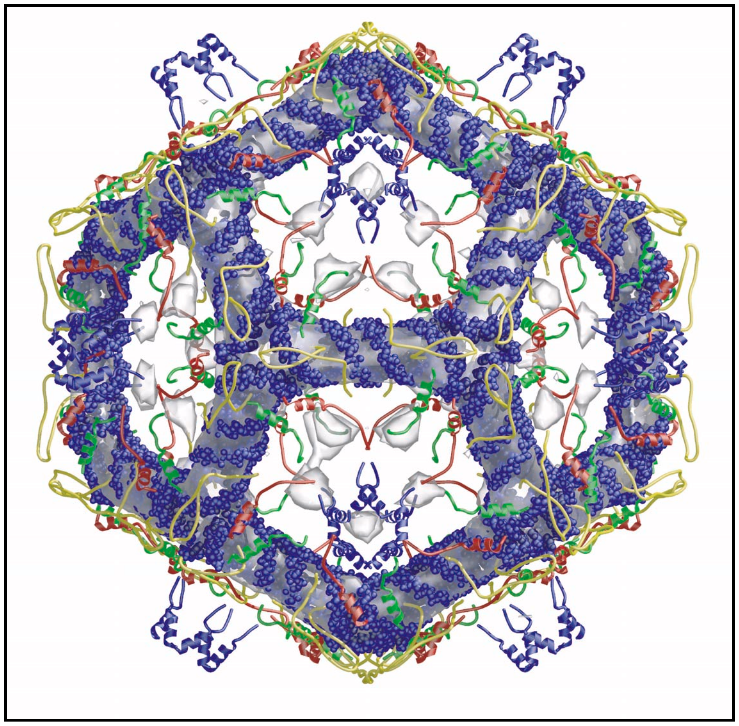

FIG. 3. The RNA cage (blue cpk model) of the crystal structure superimposed with the difference map (semitransparent, contoured at 4.2 standard deviation above background) decorated with the gamma polypeptides from three subunits (blue, red, and green, respectively) and the N-terminal segments from the A subunits (yellow).

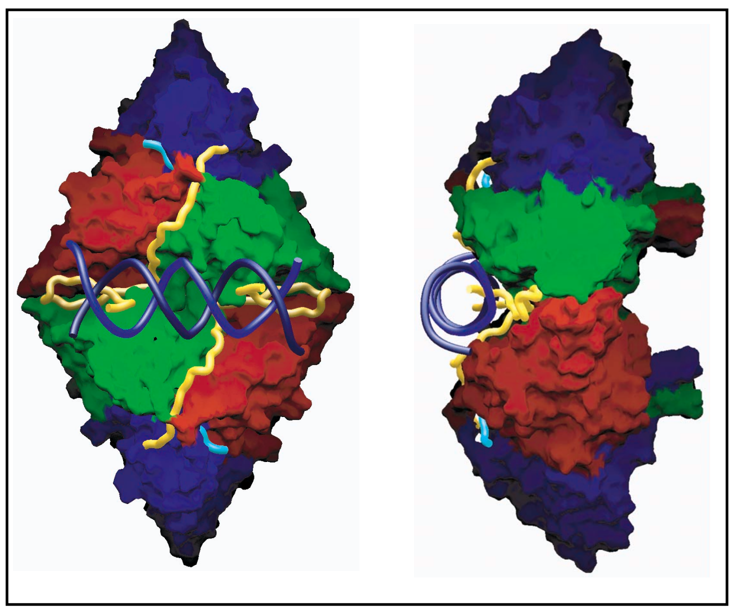

FIG. 4. Left: the interior of the capsid showing the molecular surfaces of icosahedral asymmetric units related by the 2-fold axis. Right: 90° from the left panel. The RNA duplex is shown as blue coils. The N-termini (residues 7–54) of subunits A are yellow. Only the N-terminal portions of the 8-residue C-terminal segments (cyan) are seen.

atom of Arg-9 makes H-bonds with N6 and N7 at the Hoogsteen edge of adenine-17. Lys-61 of subunit C and Arg-370 of subunits B and C are also engaged in electrostatic interactions with the ordered RNA (Fig. 2).

The ordered portion of PaV RNA is in close association with the inner surface of the protein shell that shows remarkable complementarity to the RNA in both shape and electrostatic environment. The crystal structure and cryoEM reconstruction of PaV have significantly expanded our understanding of the assembly and stability of virus particles, and have made it possible to build an atomic model for the whole viral genome packaged in the capsid and to predict steps and intermediates in the assembly pathway.

We thank the staff at the Advanced Photon Source for help in data collection and V.S. Reddy for stimulating discussion. This work was supported by NIH grants to LAB and JEJ. MY was supported by grants from the NIAID, the Gustavus and Louise Pfeiffer Research Foundation, and the Donald E. and Delia B. Baxter Foundation.

References

[1] K.N. Johnson, J.-L. Zeddam, and L.A. Ball, J. Virol. **74**, 5123-5132 (2000).

[2] P.J. Kraulis, J. Appl. Crystallogr. **24**, 946-950 (1991).

[3] Esnouf, R.M., J. Mol. Graph. **15**, 132-134 (1997).

[4] E.A. Merritt and M.E.P. Murphy, Acta Crystallogr. **D50**, 869-873 (1994).

[5] A. Nicholls, K.A. Sharp and B. Honig, Proteins **11**, 281-296 (1991).

L. Tang,[1] K. N. Johnson,[3] L. A. Ball,[3] T. Lin,[1] M. Yeager,[1,2,4] J. E. Johnson[1]

[1] Department of Molecular Biology and [2] Department of Cell Biology, The Scripps Research Institute, La Jolla, CA, U.S.A.
[3] Department of Microbiology, University of Alabama at Birmingham, Birmingham, AL, U.S.A.
[4] Division of Cardiovascular Diseases, Scripps Clinic, La Jolla, CA, U.S.A.